Epidemiol. Infect. (1996), **117**, 443–455

Preventing epidemics in a community of households

R. HALL* AND N. G. BECKER

School of Statistical Science, La Trobe University, Bundoora Vic 3083, Australia

(*Accepted 7 July 1996*)

SUMMARY

The occurrence of epidemics of vaccine-preventable diseases, and the immunization coverage required to prevent them, is affected by the presence of households and heterogeneity in the community. We consider a community where individuals live in households and are of different types, according to infectivity and/or susceptibility to infection. We describe a method for computing the critical immunization coverage to prevent epidemics in such communities and discuss the effectiveness of immunization strategies. In a heterogeneous community where individuals live in households several immunization strategies are possible and we examine strategies targeting households, randomly selected individuals, or groups with highly intense transmission, such as school children. We compare estimates of the critical immunization coverage if we assume that disease is spread solely by random mixing with estimates which result if we assume the effects of the household structure. Estimates made under these two sets of assumptions differ. The results provide insights into the community effects of vaccination, and the household structure of the community should be taken into account when designing immunization policies.

INTRODUCTION

Immunization is one of the most effective means of preventing infectious disease, but epidemics of vaccine-preventable diseases continue to occur and to cause significant mortality and morbidity, even when high levels of vaccine coverage have been achieved [1]. The prevention of epidemics of disease is a major objective of immunization programmes.

The potential of a communicable disease to cause epidemics depends on several factors; inherent characteristics of the disease, the community within which it occurs and the effectiveness of control measures. During the rising phase of an epidemic each person infected with the disease is, on average, replaced by more than one new case, and this level of replacement constitutes an epidemic threshold. If each infected person is replaced on average by less than one case, the number of cases dwindles away, and the disease will eventually disappear. If replacement occurs at a level less than one from the introductory case, then epidemics cannot occur. Replacement of cases is described by the reproduction number of a disease (R), which is the average number of secondary cases occurring as a result of transmission from a single infected person during his or her infectious period.

An immunization programme reduces the number of susceptible individuals in the community and, since transmission of disease occurs only to susceptible individuals, immunization affects the ability of infectives to replace themselves. There is a critical immunization coverage above which each infected person is replaced on average by less than one new

* Author for correspondence.

case, which removes the potential of the disease to cause epidemics. After a period, a successful immunization programme results in little or no infection with the target disease, with the consequence that individuals can acquire immunity only by vaccination. An immunization programme must therefore be capable of preventing the occurrence of epidemics in a population which would otherwise be completely susceptible.

The replacement of cases in the particular situation where all members of the community are susceptible is described by the basic reproduction number (R_0). Under the assumption that the community consists of individuals randomly mixing with one another, the minimum vaccination coverage needed to prevent epidemics is related to the basic reproduction number by the formula $v^* = 1 - 1/R_0$, where v^* is the critical vaccination coverage. This relationship was first recognized by Smith [2, fig. 3] and derived by Becker [3] and Dietz [4].

This concept of the reproduction number has offered insights into the spread of epidemic diseases and the immunization coverage required to prevent them [5, 6]. However, in this form, it is based on the idea that the community is composed of randomly mixing individuals. In reality, the community is made up of households, with higher levels of disease transmission within households and lower levels of transmission between households. The spread of a disease is modified by this community structure, and the aim of this paper is to explore the effects of community structures on the transmission of disease and on immunization programmes. We use an approach based on the approximation of the early stages of epidemic transmission by a branching process [7, 8] to examine these effects.

We show that in a community of households the replacement number of cases of an infectious disease attributable to a single initial infective depends on the household structure and increases with increasing mean and variance of the number of susceptible individuals per household (μ_S and σ_S^2 respectively). We consider the effects of different levels of transmission within households and the effects of individuals with different infectivity and/or susceptibility. We examine the design and impact of immunization strategies in a community of households it is possible to design several different immunization strategies with the same community coverage, but with different effects on the parameters μ_S and σ_S^2. For example, for the same overall community coverage, immunization may be accepted by all members of some households, or some members of all households. These different strategies have different impacts on the transmission of infection, both between and within households, and require different immunization coverages to prevent epidemics. We consider estimation of the critical immunization coverage in Section 7. Estimates of the immunization coverage required to prevent epidemics which have been determined from considering the community to be made up of homogeneously mixing individuals do not account for the impact of household structure and as a consequence differ from them. These differences may be substantial and estimates based on assumptions of uniform mixing may therefore be misleading.

THE SPREAD OF INFECTION IN A COMMUNITY OF HOUSEHOLDS

Infection in a community of households is transmitted both between and within households. We assume that infection is transmitted in the community by infected individuals' making 'adequate contact' with other individuals, that is, contact just sufficient to transmit infection to a susceptible individual. Infection is introduced into households by infected individuals who have acquired infection in the community, and transmission of disease then occurs within households with infection of some or all susceptible household members. We assume that there is no variation between households of within-household transmissibility, so that the probability of transmission to susceptibles after introduction of disease into a household is the same for all households.

Infection is thus spread from household to household and we can define a reproduction number for infected households (R_H). The reproduction numbers for individual infections and for infected households are both epidemic threshold parameters and either may be used to determine a critical immunization coverage. However, for a community of households, the reproduction number for infected households is simpler mathematically [7], and is used here.

The reproduction number for infected households is the product of the mean number of cases in a typical infected household (ν) and the mean number of adequate contacts made in the community by each of these cases with individuals from other households (b_C), so that $R_H = b_C \nu$ [9]. The mean household

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

outbreak size depends on the size of infected households and the extent of within-household transmission to other susceptibles. We find

$$R_H = \frac{b_C}{\mu_N} E(SC_S), \quad (1)$$

where μ_N is the mean size of households and $E(SC_S)$ is the mean of the product of S, the number of susceptibles in the household of an individual randomly chosen from the community, and C_S, the eventual number of cases in that household if one of its members is infected from outside [8]. A derivation of this result is given in the Appendix.

The term S appears in this relation because the probability of infection of a household is dependent on the number of susceptible individuals in the household. After the introduction of infection, the eventual number of cases (C_S) within a household then depends on the number of susceptible individuals in the household and the extent of disease transmission to them. Transmission within infected households may vary from complete, with infection of all susceptible household members, to none, with only the introductory case being infected, so that for a household with S susceptible members, C_S may take the values $1, 2, \ldots, S$ according to a probability process.

We can more easily illustrate the effect of households on the spread of disease by assuming that transmission within households is complete, so that $C_S = S$, and for diseases such as measles and pertussis complete infection of households is a good approximation [1, 10]. However, the effects hold more generally, and we discuss the situation where within-household transmission is less than complete below. Where there is complete transmission, expression (1) becomes

$$R_H = \frac{b_C}{\mu_N} E(S^2) = \frac{b_C}{\mu_N}(\mu_S^2 + \sigma_S^2), \quad (2)$$

where μ_S and σ_S are the mean and standard deviation respectively of the number of susceptible individuals in a randomly chosen household. The term $E(S^2)$ in this expression arises from the double effect of the number of susceptibles in a household on transmission, determining both the probability of household infection and the size of the subsequent household outbreak.

A basic reproduction number for infected households

The basic reproduction number for infected households (R_{H0}) obtains when all members of the community are susceptible. In this situation the number of susceptibles within a household is equal to the household size ($\mu_S = \mu_N$ and $\sigma_S = \sigma_N$) and, where there is complete transmission within households,

$$R_{H0} = b_C \mu_N \left[1 + \left(\frac{\sigma_N}{\mu_N}\right)^2\right]. \quad (3)$$

This relationship provides insight into the spread of disease in the community and permits analysis of the effects of immunization strategies. As with a uniformly mixing population, immunization in a community of households reduces the number of susceptible individuals, and to prevent epidemics, an immunization programme must be able to reduce the replacement of infectives to less than one when vaccination is the only source of immunity. Equivalently, the replacement of infected persons must be reduced so that R_H is less than one.

If all households are the same size ($\sigma_N = 0$), the basic reproduction number is linearly dependent on the mean household size. This is because transmission occurs from individual to individual and infection of any of the μ_N susceptible individuals in a household results in complete infection of that household. The effect of the variance of household size on R_{H0} arises from the double effect of household size on transmission. Since all household members are susceptible ($S = N$) and $E(N^2) = \mu_N^2 + \sigma_N^2$, if the mean household size is held constant the variance increases with the term $E(S^2)$ in expression (2). It is of interest to note that this effect is similar to the effect due to the variance of the rate of sexual partner change on the spread of sexually transmissible diseases. This similarity, for such different models, arises because both models have incorporated some heterogeneity in the potential to infect [17, equation (11.10)].

We illustrate the effects of the distribution of household sizes on the basic reproduction number in Figure 1. In this example we set $b_C = 1$ and compute R_{H0} for varying μ_N and σ_N. For a given mean household size R_{H0} increases with increasing values of σ_N, particularly when the mean household size is small. This means that a few large households, such as schools or child care facilities, may have a large impact on the transmission of communicable disease.

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

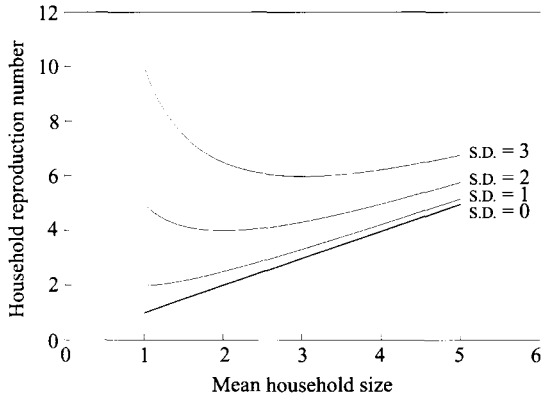

Fig. 1. R_{H0} and the mean and standard deviation of household size (S.D.). R_{H0} increases with increasing mean and standard deviation of household size.

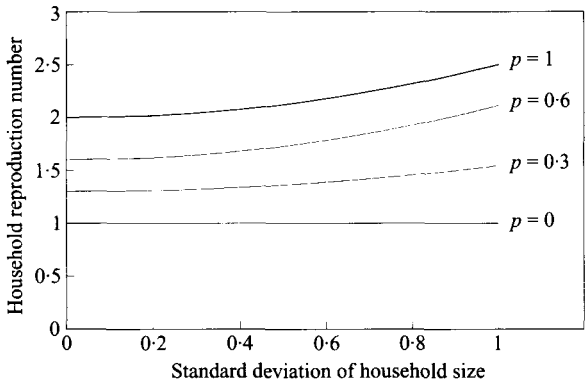

Fig. 2. R_{H0} and standard deviation of household size when there is less than complete transmission of infection within households. An example for households of sizes 1, 2 and 3. The probability of transmission from a given infected person to a given susceptible member within households is set at $p = \{0, 0{\cdot}3, 0{\cdot}6, 1\}$. R_{H0} increases with increasing standard deviation of household size for $p > 0$.

INCOMPLETE TRANSMISSION WITHIN HOUSEHOLDS

For many diseases there is less than complete transmission of infection within households. In this situation, computation of the eventual number of cases in each household requires the specification of the various probabilities of within-household disease transmission from individual to individual. As the household size increases the expression for the eventual number of cases becomes complicated, but for smaller households simple expressions may be written [11, 12].

We illustrate the effect of incomplete transmission of infection within households for the case where households consist of only one, two or three members each. If all individuals are susceptible, we may write expression (1) with an explicit term for $E(SC_S)$ (see Appendix). Then

$$R_{H0} = \frac{b_C}{\mu_N}[h_1 + 2h_2(1+p) + 3h_3(1+2p+2p^2-2p^3)], \quad (4)$$

where h_i is the proportion of households with i members, and p is the probability of disease transmission from a given infected person to a given susceptible household member. When $p = 1$ (complete transmission within households), this expression gives the same result as expression (3). If the household distribution and b_C are held constant, R_{H0} increases as the value of p increases.

The effects of incomplete transmission within households on the reproduction number are shown in Figure 2. In this example the community is made up of households of sizes 1, 2 and 3, with mean size 2 (so that the variance of household size lies in the range $0 \leqslant \sigma_N^2 \leqslant 1$; if all households are of size 2, $\sigma_N^2 = 0$, and if half the households are of size 1 and half of size 3, $\sigma_N^2 = 1$), we set $b_C = 1$, and compute R_{H0} for several values of p. When $p = 0$ there is no transmission from the introductory case to other household members and $R_{H0} = b_C$. When $p = 1$ there is complete transmission to all household members. For non-zero values of p the dependence of R_{H0} on the mean and variance of household size is maintained. The reproduction number increases with increasing variance of household size whenever there is within-household transmission, and this effect is greater for diseases which have high levels of transmission within households.

A HETEROGENEOUS COMMUNITY

We have considered so far the existence of only a single type of individual, with susceptibility and infectivity being the same for all individuals in the community. For control of communicable disease consideration of the existence of several types of individuals is essential, since individuals have markedly differing susceptibility and/or infectivity depending on age, school attendance, sex or other attributes. For example, there is much greater intensity of disease transmission in schools and child-care facilities [13].

If we classify individuals into several types by virtue of their differing susceptibility or infectivity we can generalize expression (1) to include the effects of these

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

differences. In a community of several types of individual, an infected person may be replaced by cases of several different types depending on the number of adequate contacts he or she makes with each of the different types. The basic reproduction number for infected households is determined by a matrix describing disease transmission between all the types. The typical element of this matrix, M_{ij}, is the mean number of infected type j individuals attributable to a single type i infective. We write

$$M_{ij} = \sum_{k=1}^{p} \frac{b_{ik}}{\mu_{N_k}} \mathrm{E}(S_k C_j), \tag{5}$$

where b_{ik} is the mean number of adequate contacts made by type i individuals with type k individuals in the community, μ_{N_k} is the average number of individuals of type k in a typical household, and $\mathrm{E}(S_k C_j)$ is the mean of the product of the number of type k susceptibles and the eventual number of cases of type j in a typical household where a type k individual has been infected from outside. Note that this expression for M_{ij} is similar to expression (1) describing R_{H0} for the situation where there is only one type in the community. The introduction of different types of individual complicates the computation of R_{H0}. Since adequate contacts are now made between different types in the community, there are several values for b_{ik}, there are several numbers of types of susceptible in each household (S_k) and there are several numbers of types eventually infected within households (C_j). If there are p types of individual, a matrix of the M_{ij} has p rows and p columns, and R_{H0} may then be computed as the largest eigenvalue of this matrix [8].

A practical example of the importance of the effects of different types is the higher rate of infection of children attending school or child-care facilities, compared with other individuals in the community. Using the formulation above, we can define children as being one of two types: they either attend school or child care, or they do not, and the mean number of adequate contacts made between children attending school is different from the number of contacts between children where at least one member of the possible pairs does not attend school. The reproduction number for infected households can then be calculated and increases as the transmissibility of disease among school children compared with other children increases, increasing the difficulty of epidemic prevention.

IMMUNIZATION STRATEGIES

The success of an immunization programme in preventing epidemics depends on its ability to reduce the replacement of infectious persons in a community to a value less than one, and thus to eliminate the epidemic potential of the target disease. The reproduction number for infected households with vaccination, R_{HV}, describes the replacement of infected households when the community is made up of both susceptible and vaccinated individuals and is computed in the same way as R_{H0}, but accounting for the effects of the immunization strategy on the number and distribution of susceptible individuals. The critical coverage to prevent epidemics is found by expressing R_{HV} in terms of $v_{\mathscr{S}}$, the proportion of individuals vaccinated under strategy $\mathscr{S}$, setting $R_{\mathrm{HV}} = 1$, and solving for $v_{\mathscr{S}}$.

Immunization strategies achieve their effects solely by reducing the number of susceptibles in the community and thus reducing the term $\mathrm{E}(S^2)$ in expression (2). In a community of households different strategies may be designed which affect the number and distribution of susceptible individuals in the community and within households in different ways. For the same community coverage overall, different strategies may have different effects on μ_S and σ_S^2 and therefore on $\mathrm{E}(S^2)$. Due to these differences, the critical immunization coverage required to prevent epidemics varies with the immunization strategy adopted. Strategies available include immunization of each member of a proportion of households, immunization of a fixed proportion of members of all households, immunization of a proportion of randomly selected individuals, or (where they occur) selection of specific types of individuals. We consider each of these strategies in turn for a community made up of individuals all of the same type and then for a community with different types of individuals.

In the following discussion we assume that the vaccine is fully effective in preventing infection and that there is complete transmission of the disease within households. In practice, vaccination may not confer complete protection, so the proportions vaccinated may require adjustment to give the proportions effectively immunized.

Strategy $\mathscr{S}_{\mathrm{HF}}$, immunization by household

If we immunize on a household basis, by selecting a proportion, v_{H}, of households and vaccinating a

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

proportion, v_F, of members of each selected household, the overall coverage in the community is given by the product $v_H v_F$. Only whole numbers of household members can be immunized, and because the household sizes are relatively small, v_F can take only certain values. In contrast, v_H is only minimally constrained in this way because the community is made up of a large number of households. For a typical household under this strategy, $E(S^2)$ in expression (2) becomes

$$E(S^2) = [v_H(1-v_F)^2 + (1-v_H)]\,E(N^2).$$

There are two extreme forms of this strategy. Under strategy $\mathscr{S}_H$ we vaccinate all members of the selected households, so that $v_F = 1$, and we reduce the number of susceptibles in the immunized households to zero. A proportion $(1-v_H)$ of all households then contains only susceptible individuals, and $E(S^2) = (1-v_H)\,E(N^2)$. By substitution into equation (2) we obtain

$$R_{HV} = (1-v_H)\,R_{H0}.$$

We set $R_{HV} = 1$ and solve for v_H to compute v_H^*, the critical proportion of households that needs to be immunized to prevent epidemics for this strategy, as

$$v_H^* = 1 - 1/R_{H0}.$$

Under strategy $\mathscr{S}_H$, contacts in the community made by infectious individuals with members of immunized households result in no further infection, while contacts with members of unimmunized households result in infection of entire households. Since immunized households are entirely immune they can sustain no within-household transmission, with the result that the relationship $v_H^* = 1 - 1/R_{H0}$ is also true when the level of within-household transmission in unimmunized households is less than complete.

The other extreme household strategy we label $\mathscr{S}_F$. Here we immunize a fraction v_F of members of all households ($v_H = 1$). This strategy reduces the number of susceptible individuals in each household, but leaves all households open to infection. However, the chance of infection of a household and transmission within households after the introduction of infection is reduced by the limited number of susceptibles.

Under strategy $\mathscr{S}_F$, $E(S^2) = (1-v_F)^2\,E(N^2)$. The critical coverage of vaccination is found, as before, by substitution and

$$R_{HV} = (1-v_F)^2\,R_{H0}.$$

By setting $R_{HV} = 1$ and solving for v_F we find

$$v_F^* = 1 - 1/\sqrt{R_{H0}}.$$

Where the community is made up of large households, so that the values that v_F may take are not so constrained by household size, $\mathscr{S}_F$ is a more efficient strategy than $\mathscr{S}_H$, since it requires a smaller proportion of the community to be immunized to prevent epidemics. It is more efficient because it exploits the double effect of partial immunization of households on transmission, the strategy not only reduces transmission between individuals in the community, and thus transmission between households, but also reduces within-household transmission.

Strategy $\mathscr{S}_I$, immunization of randomly selected individuals

Suppose individuals in the community choose to accept vaccination independently, in contrast to strategy $\mathscr{S}_{HF}$ where immunization is accepted on a household basis. Such a strategy, which we label $\mathscr{S}_I$, introduces variation in the immunized proportion of each household's members and an additional binomial variance term is required. Under this strategy,

$$E(S^2) = (1-v_I)^2\,E(N^2) + v_I(1-v_I)\,\mu_N$$

and

$$R_{HV} = (1-v_I)^2\,R_{H0} + v_I(1-v_I)\,b_C,$$

where v_I is the proportion of individuals immunized, μ_N is the mean household size and b_C is the number of adequate contacts made by an infected person during the infectious period in the community. The critical coverage required to prevent epidemics is dependent on both R_{H0} and b_C and is computed by setting $R_{HV} = 1$ and solving the resulting quadratic equation for v_I.

Comparison of strategies

Strategies may be compared by computing the critical immunization coverage of the community required to prevent epidemics under each strategy. If the community is made up of large households and only a single type of individual, $\mathscr{S}_F$ is the most efficient of the strategies discussed, requiring the lowest coverage in the community. However, since only whole numbers of individuals can be immunized the efficiency of $\mathscr{S}_F$

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

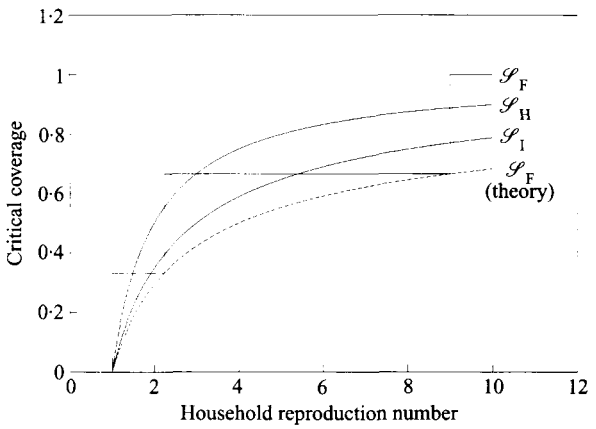

Fig. 3. Critical immunization coverage and R_{H0} for immunization strategies $\mathscr{S}_I$, $\mathscr{S}_F$ and $\mathscr{S}_H$ when all households are of size 3. Under strategy $\mathscr{S}_H$ all members of some households are immunized, under $\mathscr{S}_F$ some members of all households are immunized and under $\mathscr{S}_I$ individuals are immunized independently. Under $\mathscr{S}_F$, $v_F^* = 1 - 1/\sqrt{R_{H0}}$, the broken line (labelled $\mathscr{S}_F$ (theory)) shows the solution to this equation, the line labelled $\mathscr{S}_F$ shows the practical application of the strategy.

can be realized for only a few values of R_{H0}. The critical immunization coverages under strategies $\mathscr{S}_F$, $\mathscr{S}_H$ and $\mathscr{S}_I$ for varying values of R_{H0} in a community of households all of size 3, are illustrated in Figure 3. The critical coverages for strategies $\mathscr{S}_I$ and $\mathscr{S}_F$ are both affected by the distribution of household sizes in the community. Under $\mathscr{S}_I$, the critical coverage depends on R_{H0} and b_C (which are related to the distribution of household sizes by expression (3)), and under $\mathscr{S}_F$, v_F may take only the values 0, $\frac{1}{3}$, $\frac{2}{3}$ or 1. The pbroken line in Figure 3 shows $v_F^* = 1 - 1/\sqrt{R_{H0}}$, which sets a bound for the performance of $\mathscr{S}_F$ when households are large.

IMMUNIZATION IN CHILD CARE FACILITIES AND SCHOOLS

Children attending child care facilities and schools are often at higher risk of communicable diseases than the rest of the population and ensuring that these children are vaccinated before enrolment would appear to be an attractive strategy for preventing epidemics. We consider how this observation may be accommodated in the present formulation, where the community consists of households whose members may be of different types. Children attending child care centres or schools may be considered in two ways; these facilities may be considered as large households with high or complete infection of household members [7], or alternatively, children attending these institutions may be characterized as a separate type, defined by greater infectivity and/or susceptibility (that is, a greater number of adequate contacts is made with their own type during the infectious period) compared with other persons in the community [8].

In a community made up of individuals of types $1, 2, \ldots, p$ an additional strategy, $\mathscr{S}_j$, in which a proportion of individuals of type j is immunized independently, may be described as a variant of strategy $\mathscr{S}_I$. We assume in this discussion that the fraction of adequate contacts made by an infected person of type i with individuals of type k is in proportion to the frequency of type k in the community so that in expression (5) $b_{ik} = \mu_{N_k} b/\mu_N$, where b is the total number of potentially infectious contacts made by an infected person, μ_{N_k} is the mean number of type k individuals per household, and μ_N is the mean household size.

If transmissibility of disease among school children is higher than in the rest of the community a higher overall immunization coverage is usually required to prevent epidemics than would be the case if all children were equally susceptible and/or infectious. We give an example to illustrate the effects of differential transmission of disease among school children on the critical immunization coverage.

Suppose one third of the households in a community comprise two school children and two thirds consist of one school child and one child not attending school. We label the school children as type 1 and the others as type 2. The mean size of the households (μ_N) is 2, and it follows that the mean number of school children per household is given by $\mu_{N_1} = 1\frac{1}{3}$, and $\mu_{N_2} = \frac{2}{3}$. The transmissibility of disease among these groups of children is described by the number of adequate contacts made by an infected person of each type with the same or other types during the infectious period. We consider infectious school children to make an additional a contacts with other school children compared with contacts where at least one of the pair is not a school child. The number of infectious contacts between school children is b_{11} which we set as $b_{11} = \mu_{N_1} b/\mu_N + a$, and if the disease is less infectious among school children $a < 0$. Other permutations of contacts are

$$b_{12} = b_{22} = \frac{\mu_{N_2}}{\mu_N} b \quad \text{and} \quad b_{21} = \frac{\mu_{N_1}}{\mu_N} b,$$

and for this example we set $b = 0{\cdot}75$. This may be interpreted to mean that an infected school child makes $0{\cdot}5 + a$ contacts with other school children and

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

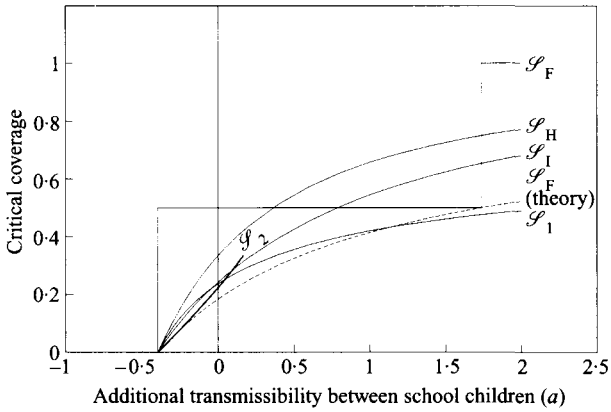

Fig. 4. Critical immunization coverages under immunization strategies $\mathscr{S}_I$, $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_F$ and $\mathscr{S}_H$ in a community of households all of size two which includes schools. Children mix in proportion to their numbers in the community, but school children make an additional a contacts with other school children. Strategies $\mathscr{S}_1$ and $\mathscr{S}_2$ consist of independent immunization of school children and children who do not attend school, respectively.

0·25 contacts with non-school children, so that the total number of contacts made by infected persons of this type is $0{\cdot}75+a$. Non-school children also make a total of 0·75 contacts with both types, and when $a=0$, in a fully susceptible community, there are equal mean numbers of infections transmitted in the community by school children and non-school children.

The number of susceptibles of each type per household (S_1 and S_2) depends on the immunization strategy, as described previously. We assume transmission within households to be complete, so that in expression (5), $\mathrm{E}(S_k C_j) = \mathrm{E}(S_k S_j)$. The reproduction number for infected households under a vaccination programme (R_{HV}) is the larger eigenvalue of the 2×2 matrix made up of the elements M_{ij}. The critical immunization coverage under each of the strategies ($v^*_{\mathscr{S}}$) is then computed by solving for $v_{\mathscr{S}}$ when $R_{HV}=1$.

We consider the immunization strategies $\mathscr{S}_1$, $\mathscr{S}_2$, $\mathscr{S}_I$, $\mathscr{S}_F$ and $\mathscr{S}_H$ and compute the critical immunization coverage for each, varying transmissibility among school children (Fig. 4). The critical immunization coverage required to prevent epidemics depends on the strategy adopted and is affected by the additional transmissibility of the disease between school children.

When $a>0$ transmissibility from school child to school child in the community is greater than transmissibility between other permutations of children. When $-0{\cdot}5 \leqslant a < 0$, transmissibility among school children is *less* than among other permutations, and if $a=-0{\cdot}5$ transmission in the community between school children does not occur at all. In the range of transmissibility from school child to school child satisfying $-0{\cdot}5 \leqslant a < -0{\cdot}4$ the disease fades out, since under these conditions $R_{H0}<1$.

Solution of the equation $v^*_F = 1-1/\sqrt{R_{H0}}$ results in lower values for v^*_F than the critical immunization coverage for other strategies. However, since only whole numbers of children can be vaccinated, v_F is restricted to the values 0, 0·5 or 1 and the efficiency of this strategy compares poorly with other strategies for most values of a. Strategy $\mathscr{S}_H$ is the least efficient of the other strategies, requiring a higher critical immunization coverage for all values of a.

We consider three strategies of independent immunization of individuals. Under strategy $\mathscr{S}_I$, individuals of both types are selected for immunization at random, under $\mathscr{S}_1$, school children are targeted and immunized independently, and strategy $\mathscr{S}_2$ consists of independent immunization of children not at school. For diseases which are highly transmissible in school settings, $\mathscr{S}_1$ is the most efficient strategy. Strategy $\mathscr{S}_2$ is capable of preventing epidemics only when transmissibility from school child to school child is low, satisfying $a \leqslant 0{\cdot}167$. At the level of transmissibility in school where $a=0{\cdot}167$, all children not attending school must be vaccinated to prevent epidemics. Since non-school children make up a third of the population, this is equivalent to a community coverage of 0·333.

For transmissibility in schools satisfying the condition $-0{\cdot}4 \leqslant a < 0{\cdot}049$, strategy $\mathscr{S}_2$ is more efficient than $\mathscr{S}_1$. This is something of a paradox for the range $0<a<0{\cdot}049$, since under these conditions transmissibility at school is higher, and there are more school children in this community than children not attending school. The explanation for this counter-intuitive result is that independent immunization of a number of non-school children reduces the mean and variance of the number of susceptibles per household to a greater extent than immunization of the same number of school children. Immunization of school children will result in immunization of two children in some households, so that in households where this occurs two vaccinations prevent within-household transmission in only one household. Immunization of two non-school children always results in immunization in two households. The household effects are so strong that, up to a point, they outweigh the effects of the additional transmission between school children.

There are two other ways in which the importance

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

of child care centres and schools for the control of infectious disease may be illustrated. First, they may be considered to be large 'households'. In a community made up of many small households (perhaps families) and a few large boarding schools, there is a large variance of household size and consequently a large reproduction number (Fig. 1). Secondly, children attending these institutions establish epidemiological links between households. Members of households with at least one child attending the same school are in effect members of a large 'household' made up of the school and the linked households [7]. Immunization of children before enrolling at school, not only reduces the number of susceptibles in schools but also breaks the links between households; whereas immunization of children at home, leaves these large epidemiological groupings intact, with the possibility of widespread transmission.

ESTIMATING THE IMMUNIZATION COVERAGE REQUIRED TO PREVENT EPIDEMICS

A practical problem for the control of vaccine-preventable disease is the estimation of the critical coverage of immunization needed to prevent epidemics. An approach is to make estimates from an analysis of the spread of disease in past epidemics. One method is based on the assumption that disease is spread by uniform mixing in the community, another on the assumption that the household structure of the community influences the spread of disease as described above. These assumptions lead to different estimates of the critical immunization coverage and the differences may be important for disease control. In this section we assume that transmission of infection within households is not necessarily complete.

If we assume that a disease spreads only by random mixing of individuals and that the household structure of the community does not influence transmission, the critical immunization coverage to prevent epidemics may be estimated from f_I, the fraction of individuals eventually infected in a previous epidemic. In a fully susceptible community, we estimate R_0 by

$$\widehat{R_0} = \frac{-\ln(1-f_I)}{(1-f_I)}$$

[14, 15] and the critical immunization coverage by

$$\widehat{v_I^{(u)}}^* = 1 - 1/\widehat{R_0}.$$

If we assume that disease transmission depends on the household structure of the community, we can estimate R_{H0} from the proportions of households of each size eventually infected during an epidemic. In a simpler case, where the community is made up of households all of the same size, the basic reproduction number for households may be estimated from the proportion of households infected during an epidemic (f_H) by

$$\widehat{R_{H0}} = \frac{-\ln(1-f_H)}{(1-f_H)}$$

[15]. We estimate the critical immunization coverage by substituting $\widehat{R_{H0}}$ into the relevant equations for the immunization strategy under consideration. For strategy $\mathscr{S}_H$, $v_H^* = 1 - 1/R_{H0}$; for strategy $\mathscr{S}_F$, $v_F^* = 1 - 1/\sqrt{R_{H0}}$, and for $\mathscr{S}_I$ we estimate the critical immunization coverage from the solution of the quadratic equation

$$(1-v_I)^2 R_{H0} + v_I(1-v_I) b_C = 1,$$

with $\widehat{b_C} = \widehat{R_{H0}}/\nu$, where ν is the mean size of household outbreaks.

Different estimates of the critical immunization coverage result from the different assumptions made about the effect of household structure of the community on disease transmission. We compare the results of these approaches by considering an example (Fig. 5). Suppose an epidemic of a vaccine-preventable disease occurs in a community where all individuals

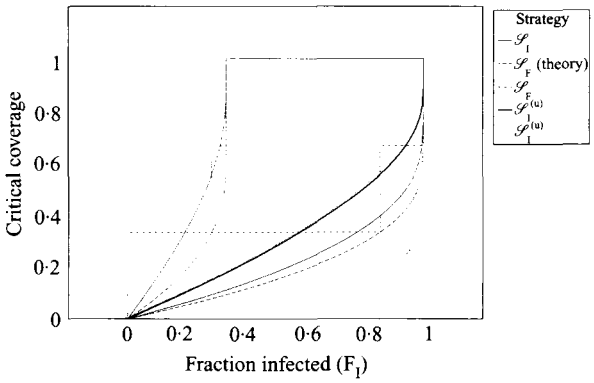

Fig. 5. Estimates of critical immunization coverages under different immunization strategies from analysis of past epidemics. In this example all households are of size 3. The mean size of household outbreaks (ν) may range from 1–3, and therefore estimates of the critical immunization coverage based on the assumption of household infection fall in ranges. Dotted lines show $\widehat{v_F^*} = 1 - 1/\sqrt{\widehat{R_{H0}}}$, but in this community v_F may take only the values 0, $\frac{1}{3}$, $\frac{2}{3}$ or 1, since only whole numbers of household members can be immunized. When ν is at a minimum, $\widehat{v_H^*} = \widehat{v_I^*}$; and when $\nu = 3$, $\widehat{v_H^*} = \widehat{v_I^{(u)}}^*$, the curves for these values have been shifted left slightly.

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

were initially susceptible and are members of households each with three members. The epidemic results in the infection of a fraction, f_I, of the community.

After the epidemic, disease surveillance data allow the computation f_I, but the fraction of households infected is not known. If infection of a household results in transmission of disease to all household members, $f_H = f_I$. However, if a disease is not transmitted from the introductory case to other members of the household, $f_H = \mu_N f_I$, until all households are infected. When all households are infected, further increase in f_I must mean that disease has been transmitted from the introductory case to other household members. Generally, the fraction of households eventually infected is related to the fraction of individuals eventually infected by $f_H = \mu_N f_I/\nu$, *where* μ_N is the mean household size and the mean size of a household outbreak satisfies $1 \leqslant \nu \leqslant \mu_N$.

As a consequence there is a range of values of $\widehat{R_{H0}}$ depending on the extent of transmission within households. The upper bound is found when there is minimum within-household transmission, and the lower bound is found when $\nu = \mu_N$. Since $\widehat{R_{H0}}$ has a range of values, the estimated critical immunization coverage also has a range of values. Note that this range arises because we have observed only f_I, if both f_I and ν are known, we obtain unique values for $\widehat{R_{H0}}$, $\widehat{v_H^*}$, $\widehat{v_F^*}$, and $\widehat{v_I^*}$.

If we assume that disease transmission results from uniform mixing of individuals in the community and that the household structure of the community is of no consequence, for each f_I we obtain a single estimate of R_0 and of the critical immunization coverage (the thick line in Fig. 5). We label the strategy arrived at under these assumptions $\mathscr{S}_I^{(u)}$.

If, however, we assume a community of households and that disease transmission occurs as described here, the mean size of the household outbreaks lies in a range $1 \leqslant \nu \leqslant 3$, depending on the characteristics of the disease. For diseases such as measles, which are highly infectious within households, ν will approach 3, for other diseases ν may be less. As discussed above, for a given value of ν, strategy $\mathscr{S}_H$ requires higher immunization coverage to prevent epidemics than $\mathscr{S}_I$. Theoretically, strategy $\mathscr{S}_F$ is the most efficient, with a lower critical immunization coverage than either $\mathscr{S}_H$ or $\mathscr{S}_I$. However, because v_F in this community may take only the values 0, $\frac{1}{3}$, $\frac{2}{3}$ or 1, the greater efficiency of $\mathscr{S}_F$ is realized only under some conditions.

An important comparison is between strategies $\mathscr{S}_I^{(u)}$ and $\mathscr{S}_I$. Under both strategies individuals are immunized independently, but $\mathscr{S}_I$ assumes that disease transmission is dependent on household structure, whereas $\mathscr{S}_I^{(u)}$ does not. For a given f_I, we find a single estimate of $v_I^{(u)*}$ and a range of estimates of v_I^* depending on the extent of within-household transmission. For a disease with a high level of transmission within households the estimate of $v_I^{(u)*}$ is greater than the estimate of v_I^*. This result means that estimates of the critical immunization coverage based on assumptions of uniform mixing are overestimates of the coverage required to prevent epidemics of diseases such as measles with intense within-household transmission.

DISCUSSION

Immunization has been one of the most successful interventions yet devised for the prevention of disease and the eradication of smallpox and the proposed eradication of polio show the potential of vaccination as a disease control strategy. However, control of highly infectious diseases such as measles has proved to be much more difficult. Models of disease transmission can play an important role in the understanding of communicable disease epidemiology and may suggest reasons for the success or failure of immunization programmes. Such models have a practical use in the planning of immunization strategies, an example being the recent British measles campaign, where all children were offered measles-containing vaccine after a model had forecast a major epidemic [16].

A limitation of many models of communicable disease transmission to date has been an assumption that the community is made up of homogeneous individuals mixing randomly with each other. This is not a realistic picture of the community and the importance of community structure and heterogeneity of the community for disease transmission has been widely recognized [6], but analytical approaches have proved difficult. Differences between age groups have been summarized in the 'Who acquires infection from whom' matrix of Anderson and May [17] and results of simulations incorporating community structure have been published [18]. A model of disease transmission which clarified the relationships between community structure, the presence of different types of individual (defined by greater or lesser infectivity and/or susceptibility), and the spread of infectious

disease would be of value in the design of disease control programmes. For public health planners, quantification of predicted outcomes under different conditions would permit rational choices between strategies.

The approach we have adopted assumes that a disease, after introduction into a household, is transmitted to some or all other susceptible household members. Each infected member then makes a number of contacts in the community, transmitting diseases to susceptible individuals, who introduce infection to their own households. We make the assumption that individuals live in well-defined households and that transmissibility between any infectious susceptible pair within a household is the same for all households. In practice, however, this may not be true and transmission of disease may vary between 'household-like' settings, such as families at home, schools, or workplaces. With these assumptions, the reproduction number for infected households is found to depend on the mean and variance of the number of susceptibles per household. The household structure of a community thus becomes a major consideration for planning disease control programmes. The mean household size in a community made up of many small and a few large households may be small, but the wide range of sizes may result in a large variance. One of the reasons for the impact of schools and child care facilities on the spread of infectious disease is that they may be considered as large 'households' and add greatly to the variance, if not the mean, of household size.

In a community of households, as distinct from a community of randomly mixing homogeneous individuals, several immunization strategies are possible. Immunization in a community of households reduces the number of susceptibles and changes the mean and variance of the number of susceptibles per household. The extent of these changes depends on the immunization strategy adopted and results in differences in efficiency for disease control between strategies. For highly infectious diseases, where disease elimination goals require highly efficient immunization strategies, a greater understanding of the role of households in transmission processes may help in control.

Immunization programmes generally target all children in a community and coverage is usually measured as the proportion of all individuals who have been vaccinated. The distribution of vaccinees in households is usually not known. Since the differences in effectiveness between immunization strategies in a community of households are so marked, data on immunization coverage by household are important to indicate which immunization strategy has, in effect, been adopted. It is possible that parents accept vaccination of children in a family on an all-or-none basis, corresponding to strategy $\mathscr{S}_H$, where all members of a selection of households are immunized, and this strategy has a relatively low efficiency.

There is intense transmission of many infectious diseases in schools and child care facilities and ensuring that all children attending these facilities are immunized is a highly efficient strategy. Furthermore, implementation of school entry immunization requirements is not logistically difficult compared with other immunization strategies.

Many countries have now adopted disease control goals which include the elimination of highly infectious vaccine-preventable diseases such as measles. The achievement of very high immunization coverage levels has been proposed in order to meet these objectives. Estimates of the critical immunization coverage made under the assumption that the community structure does not affect transmission differ from estimates made assuming a community of households. For diseases which are highly infectious within households, estimates based on models with the assumption that transmission occurs solely by random mixing in a homogeneous community are higher than estimates made based on models assuming that household structure affects transmission. For the control of diseases such as measles this may mean that a lower level of immunization coverage is required than is currently considered necessary, provided that appropriate immunization strategies are used.

An understanding of the effects of heterogeneity and community structure on the impact of immunization programmes should help define such appropriate strategies. Heterogeneity and community structure are important determinants of disease transmission and have a large impact on the difficulty of disease control and are factors which need to be considered in the design of immunization programmes.

APPENDIX

Derivation of the reproduction number for infected households

We begin by finding the mean number of infected households that are generated by a typical infected household.

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

It is assumed that members of a household go out, and make contacts, independently of each other. Let b_C be the mean number of contacts an infected individual makes with individuals from other households during the course of the infectious period. Then, given the eventual number X of infected persons in the household, the mean number of infected households generated by this infected household is $b_C X$. To make this a typical household we simply take the mean over X. Let π_c be the probability that a contact (made during the early stages of the epidemic) between an infected person and an individual selected at random from the population leads to a household outbreak of c eventual cases. The mean replacement number for a typical infected household is then given by

$$b_C \mathrm{E}(X) = b_C \sum_c c\pi_c. \tag{6}$$

That this quantity gives R_H, the reproduction number for infected households is shown in (7), below.

We now show that this expression is the same as equation (1). The probability π_c depends on both the household structure of the community and the transmission of the disease within households. Let h_{ns} be the proportion of households of size n having initially s susceptible members and let μ_N denote the mean household size.

It follows that

$$\pi_c = \sum_n \sum_{s=1} \left(\frac{nh_{ns}}{\mu_N}\right)\left(\frac{s}{n}\right)\pi_{c|ns},$$

where the term nh_{ns}/μ_N is the probability that the contact is with an individual from a household of size n having (initially) s susceptible members, the term s/n is the probability that the contacted household member is susceptible and $\pi_{c|ns}$ denotes the probability that there will be c eventual cases in a household with n members in which one of the s initial susceptibles is infected from outside the household.

Therefore

$$R_H = \frac{b_C}{\mu_N} \sum_n \sum_{s=1} \sum_{c=1} sch_{ns}\pi_{c|ns} = \frac{b_C}{\mu_N} E(SC_S), \tag{7}$$

When transmission within households is always complete $\pi_{c|ns} = 1$ for $c = s$ and zero otherwise. When within-household transmission is not always complete the computation of $\pi_{c|ns}$ requires specification of the probabilities of transmission from infected to susceptible household members. As an illustration of the latter, suppose that the community has only households of size 1, 2 and 3. Assume also that initially every individual is susceptible, so that $h_{11}+h_{22}+h_{33} = 1$. If we denote the probability of infection of a given household member by an infected member of the same household by p and write q for $1-p$, we have

$$\pi_{1|11} = 1, \quad \pi_{1|22} = q, \quad \pi_{2|22} = p, \quad \pi_{1|33} = q^2,$$
$$\pi_{2|33} = 2pq^2, \quad \pi_{3|33} = p^2(1-2q),$$

under the Reed–Frost assumptions [12, table 2.4]. Substituting these into (7) gives

$$R_{H0} = \frac{b_C}{\mu_N}[h_{11} + 2h_{22}(1+p) + 3h_{33}(1+2p+2p^2-2p^3)].$$

This is equation (4), where we have written h_i for h_{ii}.

Types of infected households

It is not clear that equation (6) gives the reproduction number for infected households, because infected households are not homogeneous so that the mean replacement number may not be the quantity that determines whether epidemics are possible. The potential to generate other infected households depends on the number of eventual cases in the household. To account for this heterogeneity we introduce types of infected households. An infected household is said to be of type i if there i eventual cases in the household.

Let $m_{ij}^{(H)}$ denote the mean number of infected households of type j generated by the i infectives of a household of type i, at the beginning of the epidemic when the depletion of susceptible households is negligible. The early proliferation of infected households is described by a multi-type branching process with mean matrix given by

$$(m_{ij}^{(H)}) = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ N \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & N \end{array} \\ \begin{pmatrix} b_C\pi_1 & b_C\pi_2 & \cdots & b_C\pi_N \\ 2b_C\pi_1 & 2b_C\pi_2 & \cdots & 2b_C\pi_N \\ \vdots & \vdots & \ddots & \vdots \\ Nb_C\pi_1 & Nb_C\pi_2 & \cdots & Nb_C\pi_N \end{pmatrix} \end{array}.$$

The basic reproduction number is the largest eigenvalue of this matrix. Because the rows of this matrix are proportional to each other makes it easy to show that the largest eigenvalue is

$$R_H = b_C \sum_{c=1}^{N} c\pi_c.$$

https://doi.org/10.1017/S0950268800059100 Published online by Cambridge University Press

The simple structure of the mean matrix is what causes the mean given in (6) to be identical with the largest eigenvalue of the above mean matrix.

REFERENCES

1. Cutts FT, Markowitz LE. Successes and failures in measles control. J Infect Dis 1994; **170** (Suppl 1): S32–41.
2. Smith CG. Factors in the transmission of virus infections from animals to man. Scient Basis Med Ann Rev 1964: 125–50.
3. Becker NG. The use of mathematical models in determining vaccination policies. Bull Int Statist Inst 1975; **46**: 478–90.
4. Dietz K. Transmission and control of arbovirus diseases. In: Ludwig D, Cooke KL, eds. Epidemiology. Philadelphia: Society for Industrial and Applied Mathematics 1975: 104–21.
5. Anderson RM, May RM. Vaccination and herd immunity to infectious diseases. Nature 1985; **318**: 323–9.
6. Fine PEM. Herd immunity: history, theory, practice. Epidemiol Rev 1993; **15**: 265–302.
7. Becker NG, Dietz K. The effect of the household distribution on transmission and control of highly infectious diseases. Math Biosci 1995; **127**: 207–19.
8. Becker NG, Hall R. Immunization levels for preventing epidemics in a community of households made up of individuals of different types. Math Biosci. In press.
9. Bartoszyński R. On a certain model of an epidemic. Applicationes Mathematicae 1972; XIII: 139–51.
10. Fine PEM, Clarkson JA, Miller E. The efficacy of pertussis vaccines under conditions of household exposure. Int J Epidemiol 1988; **17**: 635–42.
11. Bailey NTJ. Mathematical theory of infectious diseases. London: Griffin, 1975.
12. Becker NG. Analysis of infectious disease data. London: Chapman-Hall, 1989.
13. Markowitz LE, Preblud SR, Orenstein WA, et al. Patterns of transmission in measles outbreaks in the United States, 1985–1986. New Engl J Med 1989; **320**: 75–81.
14. Hethcote HW, van Ark JW. Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs. Math Biosci 1987; **84**: 85–118.
15. Becker NG. Estimation of parameters relevant for vaccination strategies. Bull Int Statist Inst 1995; Book 2: 1279–89.
16. Babad HR, Nokes DJ, Gay NJ, Miller E, Morgan-Capner P, Anderson RM. Predicting the impact of measles vaccination in England and Wales: model validation and analysis of policy options. Epidemiol Infect 1995; **114**: 319–44.
17. Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press, 1991.
18. Sattenspiel L. Epidemics in nonrandomly mixing populations: a simulation. Am J Phys Anthropol 1987; **73**: 251–65.